

# Adaptive Visual Servoing in the Presence of Intrinsic Calibration Uncertainty\*

J. Chen<sup>†</sup>, A. Behal<sup>†</sup>, D. M. Dawson<sup>†</sup>, and W. E. Dixon<sup>‡</sup>

<sup>†</sup>Department of Electrical & Computer Engineering, Clemson University, Clemson, SC 29634-0915

<sup>‡</sup>Eng. Science and Tech. Div. - Robotics, Oak Ridge National Lab., P.O. Box 2008, Oak Ridge, TN 37831-6305

E-mail: jianc, abehal, ddawson@ces.clemson.edu; dixonwe@ornl.gov

**Abstract**— In this paper, we design an adaptive kinematic controller that asymptotically regulates a robot end-effector to a desired position and orientation under visual feedback (of points located on a fixed reference frame) from a camera mounted on the end-effector. This task is accomplished despite lack of depth measurements as well as uncertainty in the camera intrinsic parameter matrix.

## I. INTRODUCTION

During the last 5 years, 2.5D visual servo controllers have become very popular given their amenability to control analysis as well as the robustness that these approaches bring to robot end-effector tracking and regulation applications. Roughly speaking, 2.5D visual servoing involves obtaining information from the 3D task-space (either through a given 3D model or more interestingly through a projective Euclidean reconstruction) to regulate the rotation and depth-related error systems while simultaneously utilizing information from the 2D image-space to regulate the planar translation related error systems. Drawing upon the aforementioned novel idea of fusing 2D and 3D information, various kinematic control strategies were designed by Malis and Chaumette [1], [2], [14], and [15] for the camera-in-hand problem. More recently, Corke and Hutchinson [5] developed a new hybrid image-based visual servoing scheme in order to decouple rotation and translation components about the  $z$ -axis from the remaining degrees of freedom so as to address the problem of desirable image-space trajectories resulting in undesirable Cartesian trajectories. Most of the above controllers conjecture that a constant, best-guess estimate of the depth information can be utilized in lieu of the exact value. Motivated by the desire to actively compensate for unmeasurable depth information, Conticelli developed an adaptive kinematic controller in [4] to ensure uniformly ultimately bounded set-point regulation provided conditions on the translational velocity and the bounds on uncertain depth parameters are satisfied. In [7], Fang et al. recently developed a 2.5D visual servo controller to asymptotically regulate a manipulator end-effector by developing an adaptive update law that actively compensates for an unknown depth parameter. In [8], Fang et

al. also developed a camera-in-hand regulation controller that incorporated a nonlinear robust control structure to compensate for uncertainty in the extrinsic calibration parameters.

Most extant monocular 2.5D visual servoing controllers rely upon knowledge of the intrinsic camera calibration parameters. In a few previous results (e.g., see [8], [15]), nominal, best-guess parameters are employed. To further engage and understand robustness issues, this paper addresses the twin problems of uncertainty in the intrinsic camera calibration parameters and lack of depth measurements. Specifically, by exploiting the triangular structure of the calibration matrix, we design an adaptive control strategy via a Lyapunov-based approach. The control problem is challenging because the unknown camera calibration matrix forces us to work with the pixel homography matrix instead of the more facile Euclidean coordinates. The problem is further complicated by the fact that the projected pixel homography has more unknowns than independent equations; hence, it is impossible to separate the rotational and translational degrees of freedom. Fortunately, the use of the plane at infinity (i.e., the vanishing points) provides additional information that can be used to surmount this issue. Given this isolation of the rotation-like information that is provided by the vanishing points, it becomes possible to embed the intrinsic parameter uncertainty in the kinematic model to facilitate adaptive control techniques for a special class of multiple-input, multiple-output systems. To the best of our knowledge, this is the first result that regulates the robot end-effector to a desired position/orientation through visual servoing by actively compensating for the lack of depth measurements and uncertainty in the camera intrinsic calibration matrix with regard to the 6 degrees-of-freedom regulation problem.

This paper is organized in the following manner. In Section II, a geometric relationship is expressed between the images of interest in terms of the orientation and position of the camera located on the robot end-effector. In Section III, we describe extraction procedures for various signals embedded in the pixel information that facilitate the control design. In Section IV, the control objective is developed along with the open-loop error dynamics. In Section V, a kinematic controller is developed via a Lyapunov based approach. Concluding remarks are given in Section VI.

\*This research was supported in part U.S. DOE Office of Biological and Environmental Research (OBER) Environmental Management Sciences Program (EMSP) project ID No. 82797 at ORNL, a subcontract to ORNL by the Florida Department of Citrus, and by U.S. NSF Grant DMI-9457967, ONR Grant N00014-99-1-0589, a DOC Grant, and an ARO Automotive Center Grant.

<b>Report Documentation Page</b>			<i>Form Approved OMB No. 0704-0188</i>	
<p>Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.</p>				
1. REPORT DATE <b>2003</b>	2. REPORT TYPE	3. DATES COVERED <b>00-00-2003 to 00-00-2003</b>		
<b>4. TITLE AND SUBTITLE</b> <b>Adaptive Visual Servoing in the Presence of Intrinsic Calibration Uncertainty</b>			5a. CONTRACT NUMBER	
			5b. GRANT NUMBER	
			5c. PROGRAM ELEMENT NUMBER	
<b>6. AUTHOR(S)</b>			5d. PROJECT NUMBER	
			5e. TASK NUMBER	
			5f. WORK UNIT NUMBER	
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b> <b>Clemson University, Department of Electrical and Computer Engineering, Clemson, SC, 29634-0915</b>			8. PERFORMING ORGANIZATION REPORT NUMBER	
<b>9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b>			10. SPONSOR/MONITOR'S ACRONYM(S)	
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
<b>12. DISTRIBUTION/AVAILABILITY STATEMENT</b> <b>Approved for public release; distribution unlimited</b>				
<b>13. SUPPLEMENTARY NOTES</b> <b>Proc. of the 2003 IEEE Conference on Decision and Control, Maui, Hawaii, U.S.A., Dec. 2003, pp. 5396-5041</b>				
<b>14. ABSTRACT</b> <b>In this paper, we design an adaptive kinematic controller that asymptotically regulates a robot end-effector to a desired position and orientation under visual feedback (of points located on a fixed reference frame) from a camera mounted on the end-effector. This task is accomplished despite lack of depth measurements as well as uncertainty in the camera intrinsic parameter matrix.</b>				
<b>15. SUBJECT TERMS</b>				
<b>16. SECURITY CLASSIFICATION OF:</b>			<b>17. LIMITATION OF ABSTRACT</b> <b>Same as Report (SAR)</b>	
a. REPORT <b>unclassified</b>	b. ABSTRACT <b>unclassified</b>	c. THIS PAGE <b>unclassified</b>	18. NUMBER OF PAGES <b>6</b>	19a. NAME OF RESPONSIBLE PERSON

## II. GEOMETRIC MODEL

Consider a reference plane  $\pi$  located on an object that has four target points<sup>1</sup> (no three of which are collinear) that are denoted by  $O_i \forall i = 1, 2, 3, 4$ . Consider two orthogonal coordinate systems, denoted by  $\mathcal{F}$  and  $\mathcal{F}^*$ , where  $\mathcal{F}$  is attached to a camera that is held by the robot end-effector, and  $\mathcal{F}^*$  is a fixed coordinate system that represents the constant, desired camera position and orientation. The Euclidean coordinates of the target points on  $\pi$  can be expressed in terms of  $\mathcal{F}$  and  $\mathcal{F}^*$ , respectively, as follows

$$\begin{aligned}\bar{m}_i(t) &\triangleq [x_i(t) \quad y_i(t) \quad z_i(t)]^T \\ \bar{m}_i^* &\triangleq [x_i^* \quad y_i^* \quad z_i^*]^T\end{aligned}\quad (1)$$

under the standard assumption that the distances from the origin of the respective coordinate frames to  $\pi$  along the focal axis remain positive (i.e.,  $z_i(t), z_i^* \geq \varepsilon > 0$  where  $\varepsilon$  is an arbitrarily small positive constant). The rotation from  $\mathcal{F}$  to  $\mathcal{F}^*$  is denoted by  $R(t) \in SO(3)$ , and the translation from  $\mathcal{F}$  to  $\mathcal{F}^*$  is denoted by  $x_f(t) \in \mathbb{R}^3$  where  $x_f(t)$  is expressed in  $\mathcal{F}$ . From the geometry between the coordinate frames depicted in Fig. 1,  $\bar{m}_i^*$  can be related to  $\bar{m}_i(t)$  as follows

$$\bar{m}_i = R\bar{m}_i^* + x_f. \quad (2)$$

Also, from the geometry given in Fig. 1, the distance  $d^* \in \mathbb{R}$  from  $\mathcal{F}^*$  to  $\pi$  along the unit normal is given by

$$d^* = n^* \bar{m}_i^* \quad (3)$$

where  $n^* \in \mathbb{R}^3$  denotes the constant unit normal to  $\pi$  expressed in  $\mathcal{F}^*$ . From (3), the relationships in (2) can be expressed as follows

$$\bar{m}_i = (R + x_h n^{*T}) \bar{m}_i^* \quad (4)$$

where  $x_h(t) \in \mathbb{R}^3$  denotes a scaled translation vector that is defined as follows

$$x_h = \frac{x_f}{d^*}. \quad (5)$$

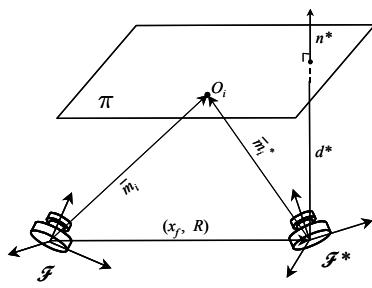


Fig. 1. Motion and structure parameters.

<sup>1</sup>It should be noted that if four coplanar target points are not available then the subsequent development can exploit the classic eight-points algorithm [14] with no four of the eight target points being coplanar.

## III. FROM GEOMETRY TO CONTROL

The relationship given in (4) is a conduit to obtaining the translational and rotational error between  $\mathcal{F}$  and  $\mathcal{F}^*$ . However, the Euclidean coordinates of the points  $\bar{m}_i(t)$  and  $\bar{m}_i^*$  cannot be directly measured; therefore, we must utilize the pixel information gathered by the camera to develop a relationship between the current and desired pixel coordinates of the points  $O_i$  in the reference plane  $\pi$ . If the intrinsic camera calibration parameters were known, standard reconstruction schemes could then be applied to obtain measurable translation and rotation error signals from which point a control algorithm could be developed to regulate  $\mathcal{F}$  to  $\mathcal{F}^*$ . However, we are interested in developing a control strategy with the additional constraint of uncertainty with regard to the intrinsic camera calibration parameters. In light of this uncertainty, the best that we can hope to do is to obtain measurable signals that mimic the rotation and translation between  $\mathcal{F}$  and  $\mathcal{F}^*$  from a decomposition of the pixel homography, and thereby, facilitate the design of kinematic control inputs that provide for asymptotic regulation of the robot end-effector. This section aims to set up the requisite signals that the subsequent control design will utilize.

### A. Euclidean and Projective Homographies

We begin by defining the normalized Euclidean coordinates  $m_i(t), m_i^* \in \mathbb{R}^3$  of  $O_i$  expressed in terms of  $\mathcal{F}$  and  $\mathcal{F}^*$ , respectively, as follows

$$m_i \triangleq \frac{\bar{m}_i}{z_i} = \left[ \begin{array}{ccc} x_i & y_i & 1 \\ z_i & z_i & 1 \end{array} \right]^T \quad (6)$$

$$m_i^* \triangleq \frac{\bar{m}_i^*}{z_i^*} = \left[ \begin{array}{ccc} x_i^* & y_i^* & 1 \\ z_i^* & z_i^* & 1 \end{array} \right]^T. \quad (7)$$

The rotation and translation between the coordinate systems can now be related in terms of the normalized coordinates as follows

$$m_i = \underbrace{\frac{z_i^*}{z_i}}_{\alpha_i} \underbrace{(R + x_h n^{*T}) m_i^*}_H \quad (8)$$

where  $\alpha_i(t) \in \mathbb{R}$  denotes the depth ratios, and  $H(t) \in \mathbb{R}^{3 \times 3}$  denotes a Euclidean homography [10].

Each target point  $O_i$  will have a projected pixel coordinate expressed in terms of  $\mathcal{F}$  (denoted by  $u_i(t), v_i(t) \in \mathbb{R}$ ) and  $\mathcal{F}^*$  (denoted by  $u_i^*, v_i^* \in \mathbb{R}$ ) that are defined as elements of  $p_i(t)$  (actual time-varying image points) and  $p_i^*$  (constant reference image points), respectively, as follows

$$p_i = [u_i \quad v_i \quad 1]^T \quad p_i^* = [u_i^* \quad v_i^* \quad 1]^T. \quad (9)$$

To calculate the Euclidean homography given in (8) from pixel information, the projected 2D pixel coordinates of the target points are related to  $m_i(t)$  and  $m_i^*$  by the following pin-hole lens model

$$p_i = A m_i \quad p_i^* = A m_i^* \quad (10)$$

where  $A \in \mathbb{R}^{3 \times 3}$  is an unknown, constant, and invertible intrinsic camera calibration matrix that is defined as follows [14]

$$A = \begin{bmatrix} a_1 & a_2 & a_4 \\ 0 & a_3 & a_5 \\ 0 & 0 & 1 \end{bmatrix} \quad (11)$$

where  $a_i \forall i = 1, 2, \dots, 5$  are scalar constants, and  $a_1, a_3 > 0$ . After substituting (8) into (10), the following relationship can be developed

$$p_i = \alpha_i \underbrace{(AHA^{-1})}_{G} p_i^* \quad (12)$$

where  $G(t) = [g_{ij}(t)] \forall i, j = 1, 2, 3 \in \mathbb{R}^{3 \times 3}$  denotes a projective homography [10].

### B. Projective Homography Decomposition

From (8) and (12), we can rewrite  $G(t)$  as follows

$$G = \bar{R} + \bar{x}_h \bar{n}^{*T} \quad (13)$$

where the signals  $\bar{R}(t) \in \mathbb{R}^{3 \times 3}$  and  $\bar{x}_h(t), \bar{n}^* \in \mathbb{R}^3$  are defined as follows

$$\bar{R} = ARA^{-1} \quad \bar{x}_h = Ax_h \quad \bar{n}^* = A^{-T}n^* . \quad (14)$$

By virtue of the definitions of (13) and (14), we can now write (12) as follows

$$p_i = \alpha_i (\bar{R} + \bar{x}_h \bar{n}^{*T}) p_i^* = \alpha_i g_{33} G_n p_i^* \quad (15)$$

where  $G_n(t) = G(t)/g_{33}(t)$  (where  $g_{33}(t) \neq 0 \forall t$  by (9) and (15)) is a normalized version of  $G(t)$ . From the relationship in (15), a set of 12 independent linear equations given by the 4 image point pairs  $(p_i(t), p_i^*)$  (3 equations for each image pair) can be used to obtain the product  $\alpha_i(t) g_{33}(t)$  and the matrix  $G_n(t)$ . Since the intrinsic camera calibration matrix  $A$  is unknown, standard techniques (e.g., [9], [19]) that rely on knowledge of the Euclidean homography  $H(t)$  cannot be applied to obtain the translational and rotational mismatch between  $\mathcal{F}$  and  $\mathcal{F}^*$ . Rather, we employ an algorithm based on vanishing points to obtain  $\bar{R}(t)$  (details available upon request). Additionally, the subsequent development requires knowledge of the depth ratios  $\alpha_i(t) \forall i = 1, 2, 3, 4$  for calculation of the translation control of the end-effector. By utilizing the relationship in (15), a nonlinear system of equations can be solved to obtain  $\alpha_i(t)$ . Details are available upon request.

## IV. CONTROL OBJECTIVE AND OPEN-LOOP ERROR SYSTEM

The objective of this paper is to develop a visual servo controller in order to ensure that the robot end-effector is in its desired final position and orientation (i.e.,  $\mathcal{F}$  is regulated to  $\mathcal{F}^*$ ). There are 6 degrees of freedom (3 translational and 3 rotational) between  $\mathcal{F}$  and  $\mathcal{F}^*$ ; hence, our control objective is satisfied if we obtain the following:  $R(t) \rightarrow I_3$  (where the notation  $I_i$  denotes an  $i \times i$  identity matrix) and  $x_f(t) \rightarrow 0$ . From (4) and (5), this amounts to  $\bar{m}_i(t) \rightarrow$

$\bar{m}_i^*$ . The rotational control objective is complicated by the fact that  $R(t)$  is not a measurable signal; however, the surrogate signal  $\bar{R}(t)$  of (14) can be utilized along with an adaptive algorithm to compensate for uncertainty in the calibration matrix  $A$ . From the second relationship of (14) and the full-rank nature of  $A$ , the translational control objective can be achieved by regulating  $\bar{x}_h(t)$  to zero. However,  $\bar{x}_h(t)$  is only available up to a scale factor that has an indeterminate sign (details available upon request). Moreover, it is desirable from a robustness point of view to servo on actual pixel information (in lieu of reconstructed information) to increase the likelihood that the object will stay in the field of view of the camera. Therefore, we state the translational control objective as<sup>2</sup>:  $p_1(t) \rightarrow p_1^*$  and  $\alpha_1(t) \rightarrow 1$ . From the definitions of (8) and (10), we obtain  $\bar{m}_1(t) \rightarrow \bar{m}_1^*$ . From the definition of (4), (5), and the fact that the control will ensure that  $R(t) \rightarrow I_3$ , we can see that  $x_f(t) \rightarrow 0$ . This chain of events ensures that the control objective is achieved (i.e.,  $\bar{m}_i(t) \rightarrow \bar{m}_i^* \forall i = 1, 2, 3, 4$ ), and hence,  $\mathcal{F}$  is regulated to  $\mathcal{F}^*$ . In light of the discussion above, we proceed to define our rotational and translation error signals whence we derive the respective open-loop error systems.

### A. Rotation Error System

Based on an angle axis parameterization of the rotation matrix  $R(t)$ , we define a rotation error-like signal  $e_\omega(t) \in \mathbb{R}^3$  as follows [15]

$$e_\omega = u\theta \quad (16)$$

where  $u(t) \in \mathbb{R}^3$  represents the unit axis of rotation (i.e., the eigenvector corresponding to the eigenvalue equal to one), and  $\theta(t)$  denotes the rotation angle about that axis. As stated in [18], the angle axis representation of (16) is not unique, in the sense that a rotation of  $-\theta(t)$  about  $-u(t)$  is equal to a rotation of  $\theta(t)$  about  $u(t)$ . A particular solution for  $\theta(t)$  and  $u(t)$  can be determined as follows [18]

$$\theta_p = \cos^{-1} \left( \frac{1}{2} (\text{tr}(R) - 1) \right) \quad [u_p]_\times = \frac{R - R^T}{2 \sin(\theta_p)} \quad (17)$$

where the notation  $\text{tr}(\cdot)$  denotes the trace of a matrix, and  $[\cdot]_\times$  denotes the following skew-symmetric matrix

$$[\zeta]_\times \triangleq \begin{bmatrix} 0 & -\zeta_3 & \zeta_2 \\ \zeta_3 & 0 & -\zeta_1 \\ -\zeta_2 & \zeta_1 & 0 \end{bmatrix} \quad \forall \zeta = \begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \end{bmatrix} . \quad (18)$$

From (17), it is clear that

$$0 \leq \theta_p(t) \leq \pi. \quad (19)$$

The constraint in (19) is consistent with the computation of  $[u(t)]_\times$  in (17) since a clockwise rotation (i.e.,  $-\pi \leq \theta(t) \leq 0$ ) is equivalent to a counterclockwise rotation (i.e.,  $0 \leq \theta(t) \leq \pi$ ) with the axis of rotation reversed.

<sup>2</sup>Any point  $O_i$  on  $\pi$  can be utilized in the subsequent development; however, to reduce the notational complexity, we have elected to select the image point  $O_1$ , and hence, the subscript 1 is utilized in lieu of  $i$  in the subsequent development of the translational controller.

Based on (17) and the functional structure of the object kinematics, the particular solutions  $\theta_p(t)$  and  $u_p(t)$  can be used in lieu of  $\theta(t)$  and  $u(t)$  without loss of generality and without confining  $\theta(t)$  to a smaller region. Since, we do not distinguish between rotations that are off by multiples of  $2\pi$ , all rotational possibilities are considered via the parameterization of (16) along with the computation of (17).

Based on the previous definitions, we can now develop the open-loop rotation error system. Since  $R(t)$  is unmeasurable and  $\bar{R}(t)$  can be calculated via a vanishing point algorithm (details available upon request), it is desirable to relate the parameterization defined in (16) to a set of measurable signals. Based on (14), we know that  $\bar{R}(t)$  is similar to  $R(t)$ , and hence,  $R(t)$  and  $\bar{R}(t)$  have the same eigenvalues. The vector  $\bar{u}(t) = [\bar{u}_1(t), \bar{u}_2(t), \bar{u}_3(t)]^T$  is the computable eigenvector of  $\bar{R}(t)$  corresponding to the eigenvalue equal to one and is related to the unmeasurable  $u(t)$  as follows

$$\bar{u} = Au. \quad (20)$$

To facilitate further development, we define an auxiliary variable  $\bar{\theta}(t)$  as follows

$$\bar{\theta} = \cos^{-1} \left( \frac{1}{2} (\text{tr}(\bar{R}) - 1) \right). \quad (21)$$

Based on the definition of  $\bar{R}(t)$  in (14) and the properties of  $\text{tr}(\cdot)$ , it is easy to see that  $\text{tr}(\bar{R}) = \text{tr}(R)$ ; hence, we can see from (17) and (21) that  $\bar{\theta}(t) = \theta(t)$ . Thus,  $\bar{u}(t)$  and  $\bar{\theta}(t)$  are measurable signals that we have now related to  $u(t)$  and  $\theta(t)$ .

The open-loop error dynamics for  $e_\omega(t)$  can be expressed as follows (details available upon request)

$$\dot{e}_\omega = -L_\omega \omega_c \quad (22)$$

where  $\omega_c(t) = [\omega_{c1}(t), \omega_{c2}(t), \omega_{c3}(t)]^T$  denotes the rotational velocity of the camera expressed in the coordinates of  $\mathcal{F}$ ,  $L_\omega(t) \in \mathbb{R}^{3 \times 3}$  is defined as

$$L_\omega = I_3 - \frac{\theta}{2} [u]_\times + \left( 1 - \frac{\text{sinc}(\theta)}{\text{sinc}^2\left(\frac{\theta}{2}\right)} \right) [u]_\times^2, \quad (23)$$

and  $u(t)$ ,  $\theta(t)$  were introduced in (17).

### B. Translation Error System

To quantify the translation mismatch between the actual and desired end-effector position, we define the translation error  $e_v(t) = [e_{v1}(t), e_{v2}(t), e_{v3}(t)]^T$  as follows

$$e_v = p_e - p_e^* \quad (24)$$

where  $p_e(t), p_e^* \in \mathbb{R}^3$  denote the extended image coordinates [15] of an image point on  $\pi$  in terms of  $\mathcal{F}$  and  $\mathcal{F}^*$ , respectively, and are defined as follows

$$p_e = [ u_1 \quad v_1 \quad \ln(z_1) ]^T \quad p_e^* = [ u_1^* \quad v_1^* \quad \ln(z_1^*) ]^T \quad (25)$$

where  $\ln(\cdot)$  denotes the natural logarithm. The signals  $u_1(t), v_1(t), u_1^*$ , and  $v_1^*$  can be measured from the two images directly. By exploiting properties of the  $\ln(\cdot)$ , the third element of  $e_v(t)$  can be expressed in terms of  $\alpha_1(t)$ . Since  $\alpha_1(t)$  can be determined,  $e_v(t)$  can be calculated. As stated earlier, the depth ratio  $\alpha_1(t)$  can be calculated; hence,  $e_v(t)$  is measurable.

We begin by observing that the extended image coordinates  $p_e(t)$  of (25) can be written as follows

$$p_e = \begin{bmatrix} a_1 & a_2 & 0 \\ 0 & a_3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{x_1}{z_1} \\ \frac{y_1}{z_1} \\ \ln(z_1) \end{bmatrix} + \begin{bmatrix} a_4 \\ a_5 \\ 0 \end{bmatrix} \quad (26)$$

where (6) and (9)-(11) were utilized. After taking the time derivative of (24) and utilizing the relationship in (26), we obtain

$$\dot{e}_v = \frac{1}{z_1} A_e \dot{m}_1 \quad (27)$$

where  $A_e(t) \in \mathbb{R}^{3 \times 3}$  is a full-rank matrix defined as follows

$$A_e = \begin{bmatrix} a_1 & a_2 & a_4 - u_1 \\ 0 & a_3 & a_5 - v_1 \\ 0 & 0 & 1 \end{bmatrix}. \quad (28)$$

After taking the time derivative of (2),  $\dot{m}_1(t)$  can be obtained as follows [15]

$$\dot{m}_1 = -v_c + [\bar{m}_1]_\times \omega_c \quad (29)$$

where  $v_c(t) = [v_{c1}(t), v_{c2}(t), v_{c3}(t)]^T$  denotes the translational velocity of the camera expressed in the coordinates of  $\mathcal{F}$ , and we have utilized the fact that  $[\omega_c]_\times = -\dot{R}R^T$ . After substituting (29) into (27), we obtain

$$\dot{e}_v = -\frac{1}{z_1} A_e v_c + \bar{Y}(\omega_c, u_1, v_1) \bar{\phi} \quad (30)$$

where  $\bar{Y}(\cdot) \bar{\phi} \in \mathbb{R}^3$  is a linear parameterization that is defined as follows

$$\bar{Y}(\cdot) \bar{\phi} = A_e [m_1]_\times \omega_c \quad (31)$$

where  $\bar{\phi} \in \mathbb{R}^m$  is the unknown constant vector of camera calibration parameters, and  $\bar{Y}(\cdot) \in \mathbb{R}^{3 \times m}$  is a measurable regression matrix. Based on (28) and the definition of  $\alpha_i(t)$  in (8), the open-loop error dynamics for  $e_v(t)$  are obtained by rewriting (30) in the following manner

$$\begin{aligned} \frac{z_1^*}{a_1} \dot{e}_{v1} &= Y_1(v_{c2}, v_{c3}, \omega_c, u_1, v_1) \phi_1 - \alpha_1 v_{c1} \\ \frac{z_1^*}{a_3} \dot{e}_{v2} &= Y_2(v_{c3}, \omega_c, u_1, v_1) \phi_2 - \alpha_1 v_{c2} \\ z_1^* \dot{e}_{v3} &= Y_3(\omega_c, u_1, v_1) \phi_3 - \alpha_1 v_{c3} \end{aligned} \quad (32)$$

where  $Y_i(\cdot) \phi_i \forall i = 1, 2, 3$  are linear parameterizations defined as follows

$$\begin{aligned} Y_1 \phi_1 &= \frac{z_1^*}{a_1} [\bar{Y} \bar{\phi}]_1 - \alpha_1 \frac{a_2}{a_1} v_{c2} + \frac{(u_1 - a_4)}{a_1} \alpha_1 v_{c3} \\ Y_2 \phi_2 &= \frac{z_1^*}{a_3} [\bar{Y} \bar{\phi}]_2 + \frac{(v_1 - a_5)}{a_3} \alpha_1 v_{c3} \\ Y_3 \phi_3 &= z_1^* [\bar{Y} \bar{\phi}]_3 \end{aligned} \quad (33)$$

where  $Y_i(\cdot) \in \mathbb{R}^{1 \times q_i}$  is a measurable regression matrix,  $\phi_i \in \mathbb{R}^{q_i}$  are unknown constant vectors, and  $[\bar{Y} \bar{\phi}]_i$  denotes the  $i^{th}$  element of the vector  $\bar{Y}(\cdot) \bar{\phi}$  of (31).

## V. CONTROL DEVELOPMENT

### A. Control Design

Before we present our camera control algorithm, a few definitions need to be introduced. We define the unknown constants  $\eta_1, \eta_2, \eta_3 \in \mathbb{R}$  as follows

$$\eta_1 = \frac{b_2}{b_3} \quad \eta_2 = b_4 \quad \eta_3 = b_5 \quad (34)$$

where  $b_i \in \mathbb{R} \forall i = 1, 2, \dots, 5$  are unknown constants and are defined to be elements of  $A^{-T}$  in the following sense

$$A^{-T} = \begin{bmatrix} b_1 & 0 & 0 \\ b_2 & b_3 & 0 \\ b_4 & b_5 & 1 \end{bmatrix}. \quad (35)$$

Consistent with the definition of  $A$  in (11) and the strict positiveness of  $a_1$  and  $a_3$  (see (11)), the following relationships hold true for each  $b_i$

$$\begin{aligned} b_1 &= \frac{1}{a_1} > 0 & b_2 &= -\frac{a_2}{a_1 a_3} & b_3 &= \frac{1}{a_3} > 0 \\ b_4 &= \frac{a_2 a_5 - a_3 a_4}{a_1 a_3} & b_5 &= -\frac{a_5}{a_3}. \end{aligned} \quad (36)$$

Based on the structure of the open-loop error system of (22), the constraints on the availability of measurable signals, and the subsequent stability analysis, the camera angular velocity control input  $\omega_c(t)$  is designed as follows

$$\begin{aligned} \omega_{c1} &= \gamma_1 \bar{\theta} \bar{u}_1 \\ \omega_{c2} &= \gamma_2 \bar{\theta} \bar{u}_2 - \hat{\eta}_1 \omega_{c1} \\ \omega_{c3} &= \gamma_3 \bar{\theta} \bar{u}_3 - \hat{\eta}_2 \omega_{c1} - \hat{\eta}_3 \omega_{c2} \end{aligned} \quad (37)$$

where  $\bar{\theta}(t), \bar{u}(t)$  have been previously defined in (20) and (21), respectively. In (37),  $\gamma_i \in \mathbb{R} \forall i = 1, 2, 3$  denote positive control gains, and  $\hat{\eta}_i(t) \in \mathbb{R} \forall i = 1, 2, 3$  denotes adaptive parameter estimates that are dynamically generated as follows

$$\dot{\hat{\eta}}_1 = -\gamma_4 \bar{\theta} \bar{u}_2 \omega_{c1} \quad \dot{\hat{\eta}}_2 = -\gamma_5 \bar{\theta} \bar{u}_3 \omega_{c1} \quad \dot{\hat{\eta}}_3 = -\gamma_6 \bar{\theta} \bar{u}_3 \omega_{c2} \quad (38)$$

where  $\gamma_i \in \mathbb{R} \forall i = 4, 5, 6$  denote positive update gains.

Motivated by the open-loop error dynamics for  $e_v(t)$  of (32) and the subsequent stability analysis, the translational control  $v_c(t)$  is designed as follows

$$\begin{aligned} v_{c1} &= \frac{1}{\alpha_1} \left( \gamma_7 e_{v1} + Y_1 \hat{\phi}_1 \right) \\ v_{c2} &= \frac{1}{\alpha_1} \left( \gamma_8 e_{v2} + Y_2 \hat{\phi}_2 \right) \\ v_{c3} &= \frac{1}{\alpha_1} \left( \gamma_9 e_{v3} + Y_3 \hat{\phi}_3 \right) \end{aligned} \quad (39)$$

where  $\gamma_i \in \mathbb{R} \forall i = 7, 8, 9$  denote positive control gains. In (39),  $\hat{\phi}_i(t) \in \mathbb{R}^{q_i} \forall i = 1, 2, 3$  denote adaptive parameter estimate vectors that are dynamically generated by the

following update laws

$$\dot{\hat{\phi}}_1 = \Gamma_1 Y_1^T e_{v1} \quad \dot{\hat{\phi}}_2 = \Gamma_2 Y_2^T e_{v2} \quad \dot{\hat{\phi}}_3 = \Gamma_3 Y_3^T e_{v3} \quad (40)$$

where  $\Gamma_i \in \mathbb{R}^{q_i \times q_i} \forall i = 1, 2, 3$  are positive diagonal gain matrices.

### B. Stability Analysis

**Theorem 1:** The adaptive update laws defined in (38) and (40) along with the control inputs designed in (37) and (39) ensure that  $e_\omega(t)$  and  $e_v(t)$  are asymptotically driven to zero in the sense that

$$\lim_{t \rightarrow \infty} \|e_\omega(t)\|, \|e_v(t)\| = 0. \quad (41)$$

**Proof:** To analyze the stability of the rotational controller, we begin by defining a non-negative function denoted by  $V_1(t) \in \mathbb{R}$  as follows

$$V_1 = \frac{1}{2} e_\omega^T e_\omega. \quad (42)$$

After taking the time derivative of (42), we obtain

$$\begin{aligned} \dot{V}_1 &= -b_1 \bar{\theta} \bar{u}_1 \omega_{c1} - b_3 \bar{\theta} \bar{u}_2 (\omega_{c2} + \eta_1 \omega_{c1}) \\ &\quad - \bar{\theta} \bar{u}_3 (\omega_{c3} + \eta_2 \omega_{c1} + \eta_3 \omega_{c2}) \end{aligned} \quad (43)$$

where we have utilized the relationships of (16), (20), (34), (35), and the fact that  $e_\omega^T L_\omega = e_\omega^T$ . The motivation for the rotational control strategy of (37) is now immediately obvious. After substituting (37) into (43), we obtain

$$\begin{aligned} \dot{V}_1 &= -b_1 \gamma_1 \bar{\theta}^2 \bar{u}_1^2 - b_3 \gamma_2 \bar{\theta}^2 \bar{u}_2^2 - \gamma_3 \bar{\theta}^2 \bar{u}_3^2 \\ &\quad - b_3 \bar{\theta} \bar{u}_2 \omega_{c1} \hat{\eta}_1 - \bar{\theta} \bar{u}_3 \omega_{c1} \hat{\eta}_2 - \bar{\theta} \bar{u}_3 \omega_{c2} \hat{\eta}_3 \end{aligned} \quad (44)$$

where the auxiliary variables  $\hat{\eta}_i(t) \in \mathbb{R} \forall i = 1, 2, 3$  are defined as

$$\tilde{\eta}_i = \eta_i - \hat{\eta}_i. \quad (45)$$

In order to take care of the last three sign indefinite terms in (44), we augment  $V_1(t)$  by defining a non-negative function denoted by  $V_\omega(t) \in \mathbb{R}$  as follows

$$V_\omega = V_1 + \frac{b_3}{2\gamma_4} \tilde{\eta}_1^2 + \frac{1}{2\gamma_5} \tilde{\eta}_2^2 + \frac{1}{2\gamma_6} \tilde{\eta}_3^2. \quad (46)$$

After taking the time derivative of (46) and utilizing (44) and the adaptive update laws of (38), we obtain

$$\dot{V}_\omega = -b_1 \gamma_1 \bar{\theta}^2 \bar{u}_1^2 - b_3 \gamma_2 \bar{\theta}^2 \bar{u}_2^2 - \gamma_3 \bar{\theta}^2 \bar{u}_3^2 \leq -\lambda_1 \theta^2 \bar{u}^T \bar{u} \quad (47)$$

where we have utilized the fact that  $\bar{\theta}(t) = \theta(t)$ , and the positive constant  $\lambda_1$  is defined as

$$\lambda_1 = \min \{b_1 \gamma_1, b_3 \gamma_2, \gamma_3\}. \quad (48)$$

Given the relationship in (20) and the fact that  $A^T A$  is positive definite, the following inequality can be developed

$$\bar{u}^T \bar{u} \geq \lambda_{\min} (A^T A) u^T u \quad (49)$$

where  $\lambda_{\min}(\cdot) \in \Re$  denotes the minimum eigenvalue of a matrix, and  $\lambda_{\min}(A^T A)$  is strictly positive. Based on (49), (47) can be rewritten as

$$\dot{V}_\omega \leq -\lambda_2 e_\omega^T e_\omega \quad (50)$$

where the strictly positive constant  $\lambda_2$  is defined as

$$\lambda_2 = \lambda_1 \lambda_{\min}(A^T A). \quad (51)$$

From (45), (46), and (50), we can now show that  $e_\omega(t) \in \mathcal{L}_2 \cap \mathcal{L}_\infty$  and that  $\tilde{\eta}_i, \dot{\tilde{\eta}}_i(t) \in \mathcal{L}_\infty \forall i = 1, 2, 3$ . From a sequential inspection of the relationships of (37) and (38), we can also show that  $\omega_c(t), \dot{\tilde{\eta}}_i(t) \in \mathcal{L}_\infty \forall i = 1, 2, 3$ . Since  $\|u(t)\| = 1$ , then  $\theta(t) \in \mathcal{L}_2 \cap \mathcal{L}_\infty$  from our preceding assertions; hence,  $L_\omega(t) \in \mathcal{L}_\infty$  from (23). From (22), we can state that  $\dot{e}_\omega(t) \in \mathcal{L}_\infty$ . Hence, we can apply Barbalat's Lemma [17] to state that  $\lim_{t \rightarrow \infty} e_\omega(t), \theta(t) = 0$ .

In order to establish the asymptotic regulation of the camera translational error, we define a non-negative function denoted by  $V_v(t) \in \Re$  as follows

$$\begin{aligned} V_v &= \frac{1}{2} \left( \frac{z_1^*}{a_1} \right) e_{v1}^2 + \frac{1}{2} \left( \frac{z_2^*}{a_2} \right) e_{v2}^2 + \frac{1}{2} z_1^* e_{v3}^2 \quad (52) \\ &+ \frac{1}{2} \tilde{\phi}_1^T \Gamma_1^{-1} \tilde{\phi}_1 + \frac{1}{2} \tilde{\phi}_2^T \Gamma_2^{-1} \tilde{\phi}_2 + \frac{1}{2} \tilde{\phi}_3^T \Gamma_3^{-1} \tilde{\phi}_3 \end{aligned}$$

where the auxiliary variables  $\tilde{\phi}_i(t) \in \Re \forall i = 1, 2, 3$  are defined as

$$\tilde{\phi}_i = \phi_i - \hat{\phi}_i. \quad (53)$$

After taking the time derivative of (52) and utilizing (32), (39), and (40), the following expression can be obtained

$$\dot{V}_v = -\gamma_7 e_{v1}^2 - \gamma_8 e_{v2}^2 - \gamma_9 e_{v3}^2 \leq 0. \quad (54)$$

Based on (52), (53), and (54), it is easy to see that  $e_v(t) \in \mathcal{L}_\infty \cap \mathcal{L}_2$  and  $\tilde{\phi}_i(t), \dot{\tilde{\phi}}_i(t) \in \mathcal{L}_\infty$ . From the definitions of (24) and (25),  $u_1(t), v_1(t), z_1(t) \in \mathcal{L}_\infty$ . From the positivity constraint placed on  $z_i^*$ , we can see that  $\alpha_i^{-1}(t) \in \mathcal{L}_\infty$ . From (39), we can see that  $v_c(t) \in \mathcal{L}_\infty$ . Given the preceding assertions, we can use (32) to show that  $\dot{e}_v(t) \in \mathcal{L}_\infty$ . Since  $e_v(t) \in \mathcal{L}_\infty \cap \mathcal{L}_2$  and  $\dot{e}_v(t) \in \mathcal{L}_\infty$ , Barbalat's Lemma [17] can be applied to state that  $\lim_{t \rightarrow \infty} e_v(t) = 0$ . Hence, we have proved the result of (41).  $\square$

## VI. CONCLUSIONS

In this paper, a kinematic controller is developed that asymptotically forces a robot end-effector to its desired position and orientation while utilizing a camera fixed on the robot end-effector for visual feedback. A homography-based algorithm that utilizes feature point recognition on a target of interest as well as points on the plane at infinity are utilized in conjunction with an adaptive camera controller in order to obtain the result. Since the controller is designed to be differentiable, robot dynamic effects can easily be incorporated via backstepping design methods [7]. Future research will target the development of adaptive visual servo controllers for the fixed camera configuration.

## REFERENCES

- [1] F. Chaumette and E. Malis, "2 1/2 D Visual Servoing: A Possible Solution to Improve Image-Based and Position-Based Visual Servoing," *Proc. of the IEEE International Conference on Robotics and Automation*, San Francisco, California, May 2000, pp. 630-635.
- [2] F. Chaumette, E. Malis, and S. Boudet, "2D 1/2 Visual Servoing with Respect to a Planar Object," *Proc. of the Workshop on New Trends in Image-Based Robot Servoing*, 1997, pp. 45-52.
- [3] J. Chen, A. Behal, D. Dawson, and Y. Fang, "2.5D Visual Servoing with a Fixed Camera," *Proc. of the American Control Conference*, Denver Colorado, June 2003, pp. 3442-3447.
- [4] F. Conticelli and B. Allotta, "Nonlinear Controllability and Stability Analysis of Adaptive Image-Based Systems," *IEEE Transactions on Robotics and Automation*, Vol. 17, No. 2, pp. 208-215 (2001).
- [5] P. I. Corke and S. A. Hutchinson, "A New Hybrid Image-Based Visual Servo Control Scheme," *Proc. of the IEEE Conference on Decision and Control*, Sydney, Australia, December 2000, pp. 2521-2527.
- [6] K. Deguchi, "Optimal Motion Control for Image-Based Visual Servoing by Decoupling Translation and Rotation," *Proc. of the Intl. Conf. on Intelligent Robots and Systems*, B.C. Canada, pp. 705-711, Oct. 1998.
- [7] Y. Fang, A. Behal, W.E. Dixon, and D. M. Dawson, "Adaptive 2.5D Visual Servoing of Kinematically Redundant Robot Manipulators," *Proc. of the IEEE Conference on Decision and Control*, Las Vegas, NV, pp. 2860-2865, December 2002.
- [8] Y. Fang, W. E. Dixon, D. M. Dawson, and J. Chen, "Robust 2.5D Visual Servoing for Robot Manipulators," *Proc. of the IEEE American Control Conference*, Denver, Colorado, June 2003, pp. 3311-3316.
- [9] O. Faugeras and F. Lustman, "Motion and structure from motion in a piecewise planar environment," *International Journal of Pattern Recognition and Artificial Intelligence*, Vol. 2, No. 3, pp. 485-508, (1988).
- [10] O. Faugeras, *Three-Dimensional Computer Vision*, The MIT Press, Cambridge, Massachusetts, 2001.
- [11] C. A. Felippa, *A Systematic Approach to the Element-Independent Corotational Dynamics of Finite Elements*, Center for Aerospace Structures Document Number CU-CAS-00-03, College of Engineering, University of Colorado, January 2000.
- [12] G. D. Hager and S. Hutchinson (guest editors), Special Section on Vision-Based Control of Robot Manipulators, *IEEE Trans. Robotics and Automation*, Vol. 12, No. 5, (1996).
- [13] E. Malis, "Contributions à la Modélisation et à la Commande en Asservissement Visuel," Ph.D. Dissertation, University of Rennes I, IRISA, France, Nov. 1998.
- [14] E. Malis and F. Chaumette, "2 1/2 D Visual Servoing with Respect to Unknown Objects Through a New Estimation Scheme of Camera Displacement," *International Journal of Computer Vision*, Vol. 37, No. 1, pp. 79-97, (2000).
- [15] E. Malis, F. Chaumette, and S. Boudet, "2 1/2 D Visual Servoing," *IEEE Transactions on Robotics and Automation*, Vol. 15, No. 2, pp. 238-250, (1999).
- [16] B. Nelson and N. Papanikopoulos (guest editors), Special Issue of Visual Servoing, *IEEE Robotics and Automation Mag.*, Vol. 5, No. 4, (1998).
- [17] J. J. E. Slotine and W. Li, *Applied Nonlinear Control*, Prentice Hall, Inc: Englewood Cliff, New Jersey, 1991.
- [18] M. W. Spong and M. Vidyasagar, *Robot Dynamic and Control*, John Wiley and Sons, Inc: New York, New York, 1989.
- [19] Z. Zhang and A. R. Hanson, "Scaled Euclidean 3D Reconstruction Based on Externally Uncalibrated Cameras," *IEEE Symp. on Computer Vision*, pp. 37-42, 1995.